DOCUMENT RESUME

ED 048 095 SP 004 648

AUTHOR Hayes, Robert B.
TITLE A Way to Evaluate and To Improve Classroom Teaching Effectiveness.
INSTITUTION Pennsylvania State Dept. of Education, Harrisburg.
PUB DATE 68
NOTE 10p.

EDRS PRICE EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS *Measurement Instruments, Reliability, *Student Attitudes, *Teacher Evaluation, Validity
IDENTIFIERS *Hayes Pupil Teacher Reaction Scale

ABSTRACT

This paper reports results of efforts over a 7-year period (1960-67) to determine if the Hayes Pupil-Teacher Reaction Scale is a reliable, valid unidimensional instrument which may be used to measure the attitude of students toward the teaching effectiveness of their teachers. Criteria used were 1) each respondent's total score describes with at least 90 percent accuracy which items were reacted to favorably or unfavorably; 2) the item marginals, or the percentage of favorable responses to each scale item can be consistently rank ordered from the lowest percentage pro through the highest percentage pro; and 3) the scale can be submitted successfully to intensity analysis to determine which score represented a dividing line between favorable and unfavorable attitudes. Test included administration to 1) 1,070 university undergraduates who selected and rated one of their best teachers and one of their worst; 2) a follow-up with 660 college sophomores; 3) 2,186 tenth graders in 14 high schools--using Cornell scalogram analysis and intensity analysis and correlation with principal's ratings of teachers, 4) replication with 1,912 sixth graders in 48 schools in seven districts. Conclusions were that the Hayes Scale, which takes only a few minutes to administer and to analyze, appears to provide a reliable, reasonably valid way to help teachers improve their teaching. It also possesses some characteristics of unidimensionality. (The instrument is included.) (JS)

ED048095

COMMONWEALTH OF PENNSYLVANIA
Department of Public Instruction
Bureau of Research

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEH REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

A Way to Evaluate and to Improve Classroom Teaching Effectiveness

Author: Dr. Robert B. Hayes

Education has been criticized because it has not developed a satisfactory yardstick to measure teaching effectiveness (Hechinger, 1961). A satisfactory measuring instrument is needed so teachers may see more clearly what they should do to improve their teaching. Barr (1948 and 1952) examined many investigations of teaching efficiency of the past several decades and found that they pay insufficient attention to "The particulars of teaching" and the relationship between teacher and student. Gage, Chatterjee and Runkel (1960) have found that sixth grade teachers will modify their teaching in the light of pupil rating. Remmers (1963) has concluded: "If 25 or more student ratings are averaged, they are as reliable as the better educational and mental tests available."

The purpose of this study was to determine if the Hayes Pupil-Teacher Reaction Scale is a reliable, valid unidimensional instrument which may be used to measure the attitude of students toward the teaching effectiveness of their teachers. A satisfactory unidimensional instrument includes these factors: (1) each respondent's total score describes with at least ninety percent accuracy which items were reacted to favorably or unfavorably; (2) the item marginals, or the percentage of favorable responses to each scale item can be consistently rank ordered from the lowest percentage pro through the highest percentage pro; and (3) the scale can be submitted successfully to intensity analysis (Stouffer et al., (1950) to determine which score represents a dividing line between favorable and unfavorable attitudes. It may also be said with 90 percent accuracy that any teacher with a higher score than another teacher was rated favorably on the same items plus a favorable rating to one or more additional items. In other words, teachers may be ranked in a consistent order in terms of degrees of effectiveness by use of an unidimensional instrument.

Several administrations and refinements in 1960 of the Hayes Scale indicated that there are certain desirable behaviors which are generally characteristic of good teachers and that these behaviors are not generally characteristic of poor teachers (Hayes). Here one thousand and seventy undergraduates at The Pennsylvania State University selected and rated one of their best and then one of their worst instructors with the following results:

P004648

Table I

Comparison of 1070 Best and 1070 Poorest Instructors

Attribute	Instructors Best	Poorest
Makes objective clear when he begins and maintains interest all the time.....	98%	10%
Is weak at stimulating thought.....	3%	65%
Explanations often not clear.....	2%	90%
Inadequate supporting materials.....	3%	88%
Provides very well for interests, needs, and experience level of students.....	96%	9%
Uses excellent examples.....	92%	14%
Instruction is very realistic and challenging.....	97%	9%
Abruptly concludes lessons, and I often wonder what I should have learned.....	1%	74%

In 1961 a follow-up study with 660 college sophomores (Hayes, 1963) produced evidence of the unidimensionality of the Hayes Scale. The item coefficients of reproducibility ranged from .87 to .96. The item marginals or percentage pro could be arranged along the following continuum: .26, .44, .55, .63, .71, .76, .83, .89, and .93. Intensity analysis (See Figure 1) could be performed and this indicated that a score of 5 on the nine items was the dividing line betwen unfavorable and favorable attitudes.

In 1965-66 the Hayes Scale was administered to some 2186 tenth grade students in 14 high schools of Pennsylvania (Hayes, Keim, and Neiman, 1966). Here Cornell scalogram analysis (Guttman, 1947) produced an average coefficient of reproducibility of .87 for the nine items comprising the Hayes Scale. The average percent for item marginals was .50 and these marginals or favorable response proportions could be consistently rank ordered along a continuum ranging from .24 to .76. Intensity analysis (Suchman and Guttman, 1947) was again successfully performed with attitude scores of one in the fall and three in the spring being the dividing points (See Figure 2) between favorable and unfavorable attitudes toward the teaching of teachers. Coefficients of consistency of .73 were obtained for 1683 students when there were three weeks between ratings and .67 for 1475 students when there were 25 weeks between ratings. These student ratings also correlated .35 with principals' ratings of the teachers. Ninety-four percent of the teachers who received the results of pupil ratings felt that pupil-teacher ratings should be widely used by high school teachers to improve their teaching.

Figure 1

INTENSITY CURVE FOR COLLEGE SOPHOMORES

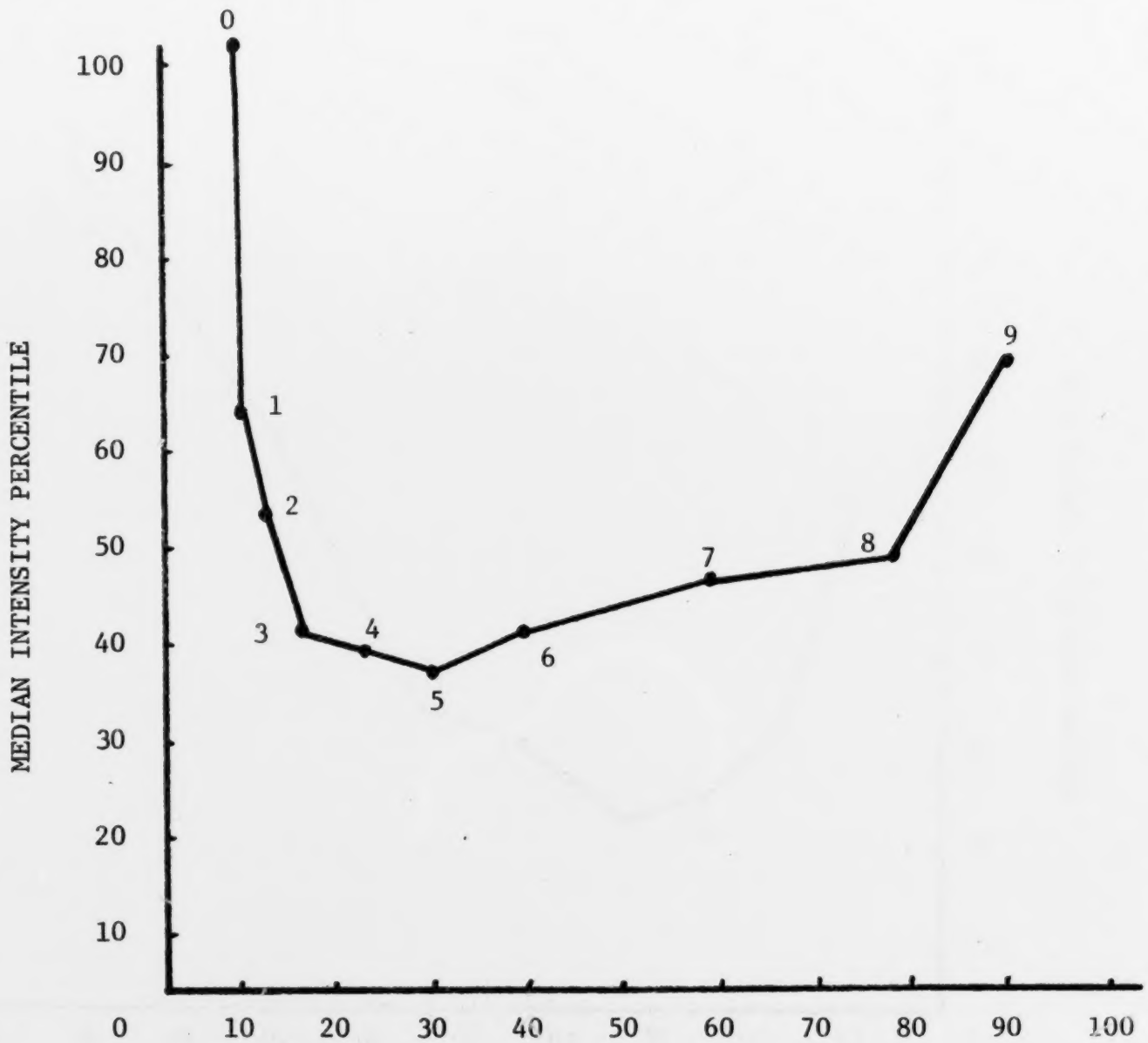
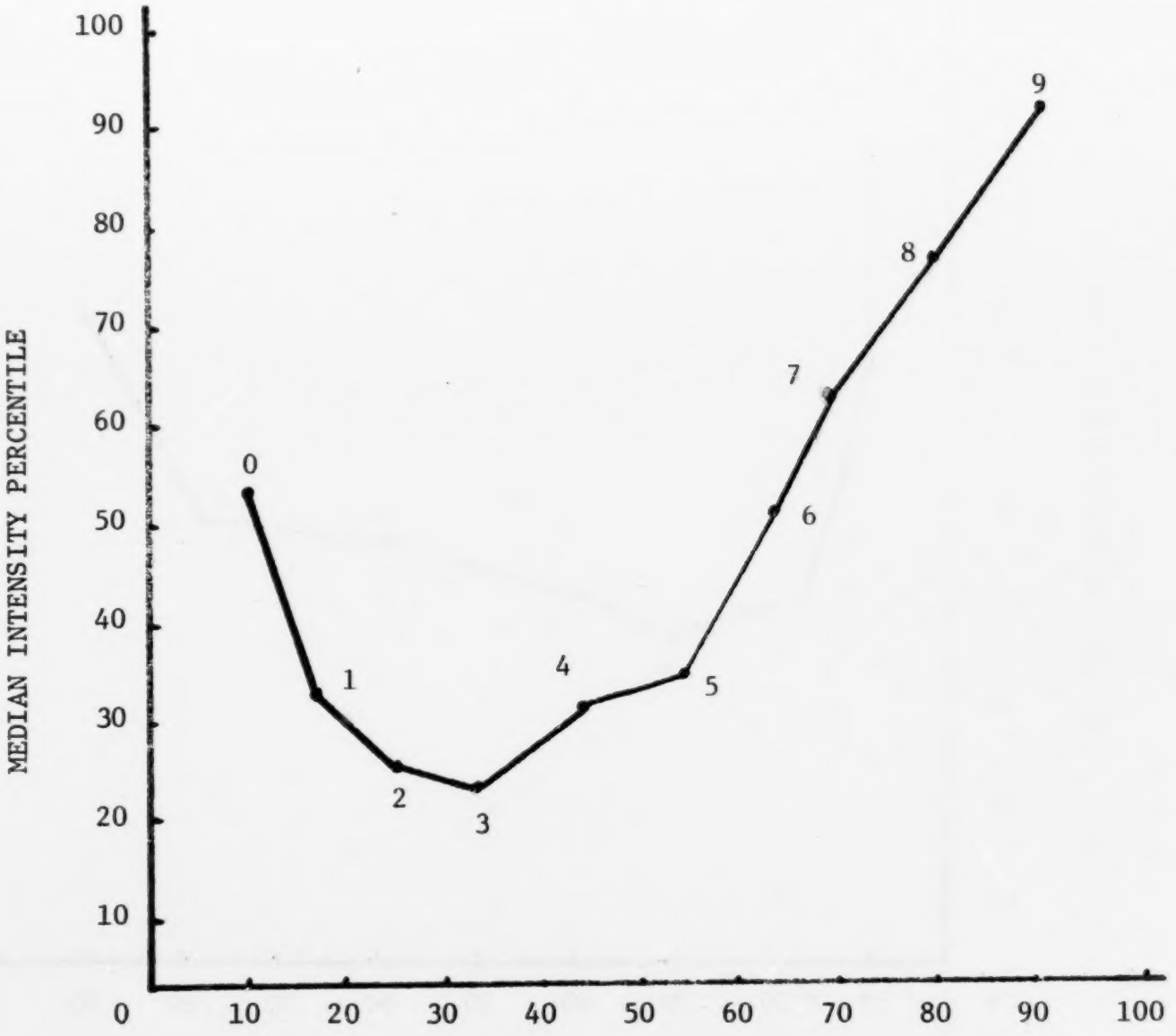


Figure 2

INTENSITY CURVE FOR TENTH GRADE SAMPLE



In 1966-67 the Hayes Scale was administered to some 1912 sixth grade pupils in 48 school buildings and 7 school districts of the beautiful Lehigh Valley of Pennsylvania (Hayes, Keim, and Neiman, 1967). Cornell scalogram analysis resulted in an average coefficient of reproducibility of .82 for the nine items of the Hayes Scale. The average for the item marginals was .50 and they could be arranged along a continuum from .26 to .80. Intensity analysis (See Figure 3) was performed satisfactorily in both fall and spring and both times any score but zero represented a favorable attitude in the minds of the pupils. Coefficients of consistency on various administrations ranged from .58 (with 21 weeks between ratings) to .85 (with 5 weeks between ratings). The pupils' ratings correlated .27 with principals' ratings of the teachers. The correlation between student attitude toward teaching of teachers (as measured by the Hayes Scale) and student attitude toward school subjects (as measured by the Remmers Scale) was .37 in the fall and .45 in the spring. Ninety percent of the teachers who received feedback on their pupil ratings felt that most of their pupils accurately rated their teaching.

In conclusion, it should be noted that the Hayes Scale takes only a few minutes to administer and to analyze and yet it appears to provide a reliable, reasonably valid way to help teachers improve their teaching. Also, the instrument possesses some characteristics of unidimensionality.

The items on the Hayes Scale are as follows:

1. This teacher makes the lesson objectives clear in the first few minutes of the class:

    a. Always
    b. Usually
    c. Sometimes
    d. Seldom or never

2. She (he) really causes you to think:

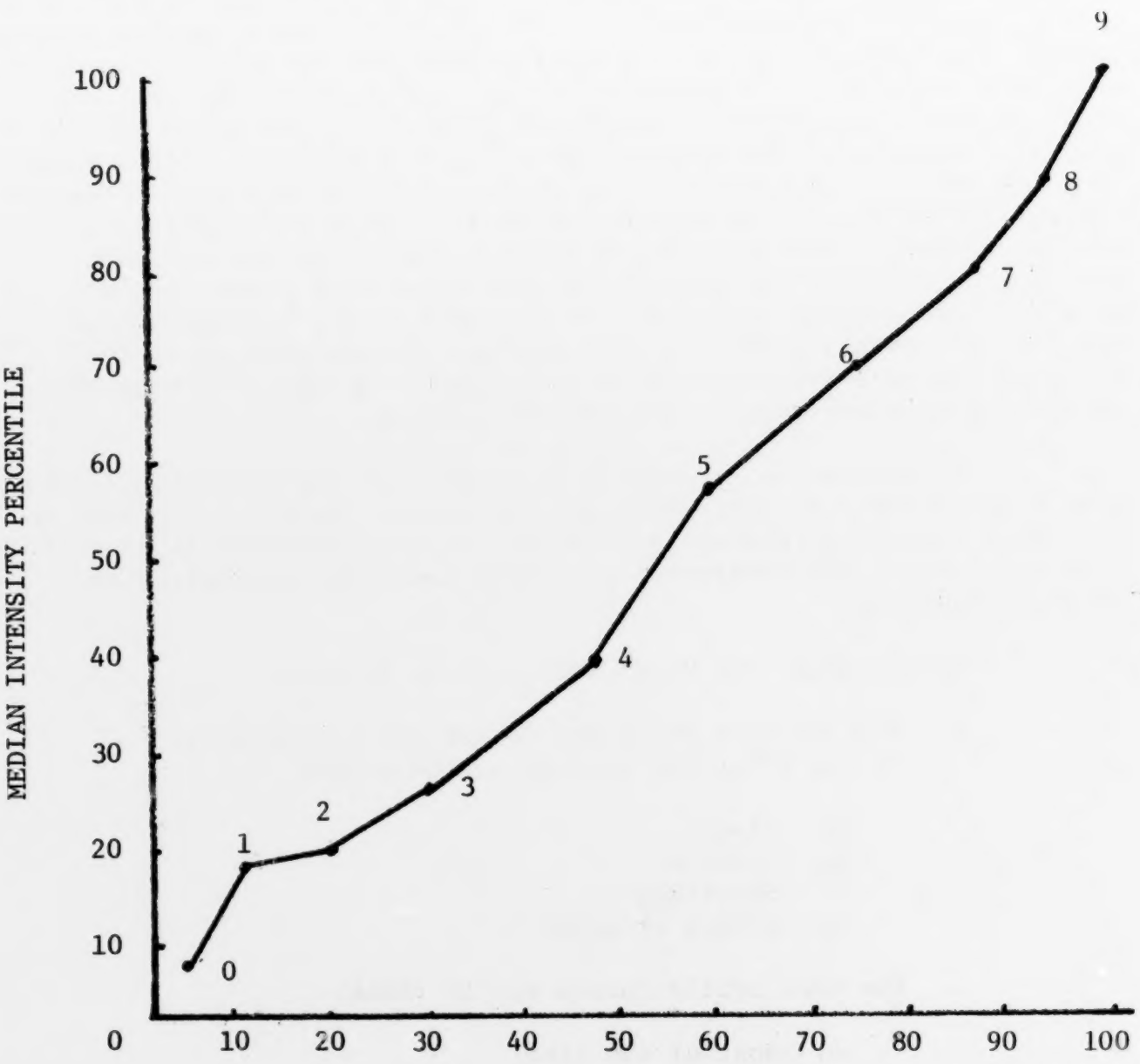
    a. Most of the time
    b. Often
    c. Sometimes
    d. Seldom or never

3. Her (his) explanations are:

    a. Extremely clear and to the point
    b. Very clear and to the point
    c. Adequate, might be better
    d. Often not clear or not to the point

Figure 3

INTENSITY CURVE FOR SIXTH GRADE SAMPLE



MIDPOINT OF CONTENT PERCENTILE

4. Her (his) lesson materials are:

    a. Outstanding
    b. Very good
    c. About average
    d. Definitely below average

5. Her (his) lessons provide very well for the needs, interests, and experience level of students:

    a. Always
    b. Usually
    c. About half the time
    d. Sometimes or seldom

6. Her (his) instruction is very realistic:

    a. Always
    b. Often
    c. Sometimes
    d. Seldom or never

7. Her (his) instruction is:

    a. Extremely challenging
    b. Very challenging
    c. Somewhat challenging
    d. Not very challenging or usually unchallenging

8. She (he) concludes lessons by:

    a. Capably emphasizing the main points
    b. Repeating the main points
    c. Abruptly stopping, but this does not bother me
    d. Abruptly stopping and I often wonder what I should have learned during the period

9. This teacher uses excellent examples to make ideas clear:

    a. Most of the time
    b. Usually
    c. About half the time
    d. Sometimes or seldom

The dichotomized scoring keys (Edwards, 1957) were:

Scoring Key Comparison

Item	1961 Penn State	1965-66 Tenth Grade	1966-67 Sixth Grade
1	a or b = 1	a or b = 1	a = 1
2	a or b = 1	a or b = 1	a or b = 1
3	a or b = 1	a or b = 1	a = 1
4	a or b = 1	a or b = 1	a or b = 1
5	a or b = 1	a = 1	a = 1
6	a or b = 1	a = 1	a = 1
7	a or b = 1	a or b = 1	a or b = 1
8	a, b or c = 1	a or b = 1	a = 1
9	a or b = 1	a = 1	a = 1

A comparison of coefficients of reproducibility and of item marginals for the three times of administration follows in Tables II and III.

Table II

Coefficients of Reproducibility

	Item	1961 Penn State	1965-66 Tenth Grade	1966-67 Sixth Grade
1.	Objectives	.95	.88	.82
2.	Thinking	.91	.86	.83
3.	Explanations	.90	.89	.83
4.	Materials	.88	.81	.87
5.	Student Needs	.91	.89	.81
6.	Realistic	.91	.81	.80
7.	Challenging	.94	.86	.84
8.	Conclusions	.96	.88	.75
9.	Examples	.87	.88	.79
	Average	.91	.86	.82

Table III

Item Marginals (% Pro Response)

Item		1961 Penn State	1965-66 Tenth Grade	1966-67 Sixth Grade
7.	(Challenging)	.26	.54	.70
2.	(Thinking)	.44	.43	.77
4.	(Materials)	.55	.64	.80
3.	(Explanations)	.63	.72	.26
5.	(Needs)	.71	.33	.40
9.	(Examples	.73	.39	.56
6.	(Realistic)	.76	.50	.39
1.	(Objectives)	.89	.72	.30
8.	(Conclusions)	.93	.24	.36
	Average	.66	.50	.50

Bibliography

1. Barr, Arvil S. "The Measurement and Prediction of Teaching Efficiency - A Summary of Investigations." Journal of Experimental Education, 16:216, June 1948.

2. Barr, Arvil S. "The Measurement of Teacher Characteristics and Predictions of Teaching Efficiency." Review of Educational Research, 22:172, June 1952.

3. Edwards, Allen L. Techniques of Attitude Scale Construction, (New York: Appleton-Century-Crofts, Inc., 1957), pp.212-13.

4. Gage, N. L., Chatterjee, B.B. and Runkel, Philip F. Equilibrium Theory and Behavior Change: An Experiment from Pupils to Teachers, (Urbana, Illinois: Bureau of Research, University of Illinois, 1960), p.90.

5. Guttman, Louis. "The Cornell Technique for Scale and Intensity Analysis." Educational and Psychological Measurement, 7:247-79, Summer 1947.

6. Hayes, Robert B. "A Measure of Student Attitude Toward Teaching Effectiveness." Unpublished D.Ed. dissertation, The Pennsylvania State University, 1961, pp.107-108.

7. Hayes, Robert B., Keim, Floyd N. and Neiman, Albert M., The Effects of Student Reactions to Teaching Methods, Project No. 5-0096, USOE Contract No. OE-6-10-018, November 1966.

8. Hayes, Robert B., Keim, Floyd N. and Neiman, Albert M., The Effects of Student Reactions to Teaching Methods, Project No. 6-2056, USOE Contract No. OEG-1-7-062056-0302, September 1967.

9. Hechinger, Fred M. "Missing Yardstick." The New York Times, Section 4:8E, July 9, 1961.

10. Remmers, H. H. "Rating Methods in Research on Teaching." Handbook of Research on Teaching (Edited by Gage, Chicago: Rand - McNally, 1963), p.367.

11. Suchman, Edward A. and Guttman, Louis. "A Solution to the Problem of Question Bias." Public Opinion Quarterly, 11:445-55, Fall 1947.

12. Stouffer, Samuel A., et al. Measurement and Prediction: Studies in Social Psychology in World War II, IV. (Princeton: Princeton University Press, 1950). pp.249-53.